A Constrained Matching Pursuit Approach to Audio Declipping

Amir Adler, Valentin Emiya, Maria G. Jafari, Michael Elad, Rémi Gribonval, Mark D. Plumbley

A CONSTRAINED MATCHING PURSUIT APPROACH TO AUDIO DECLIPPING

Amir Adler[⋆], *Valentin Emiya*[†], *Maria G. Jafari*[⋄], *Michael Elad*[⋆], *Rémi Gribonval*[†], *Mark D. Plumbley*[⋄]

[⋆] Computer Science Department, The Technion, Haifa 32000, Israel
[†] INRIA, Centre Inria Rennes - Bretagne Atlantique, 35042 Rennes Cedex, France
[⋄] Queen Mary University of London, Centre for Digital Music,
School of Electronic Engineering and Computer Science, London E1 4NS, U.K

ABSTRACT

We present a novel sparse representation based approach for the restoration of clipped audio signals. In the proposed approach, the clipped signal is decomposed into overlapping frames and the declipping problem is formulated as an inverse problem, per audio frame. This problem is further solved by a constrained matching pursuit algorithm, that exploits the sign pattern of the clipped samples and their maximal absolute value. Performance evaluation with a collection of music and speech signals demonstrate superior results compared to existing algorithms, over a wide range of clipping levels.

Index Terms— Audio, Clipping, Inpainting, Sparsity, OMP

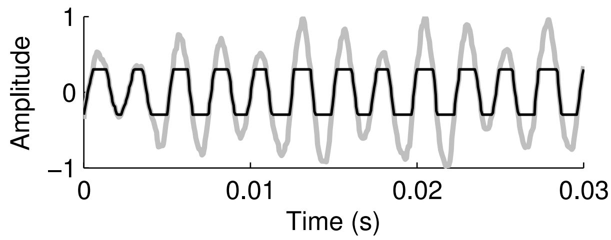

Fig. 1. A speech signal (gray) and its clipped version (black).

1. INTRODUCTION

Audio clipping is a signal degradation process in which an undistorted audio waveform is truncated whenever the maximum input range of a digital acquisition system is exceeded, as illustrated in Fig. 1. Although clipped audio signals are often encountered in telephony systems, low-cost digital audio/video recorders and other devices, restoring clipped signals has attracted substantially limited research efforts [1–4] compared to other audio restoration tasks such as click removal (see [5] for a review). In the click removal problem, samples randomly distorted by impulsive noise or small spikes (typical to old recordings or scratched CDs) are recovered. In the clipped audio case, the problem is even more challenging as the clipped samples are arranged in groups and their location is not random but rather determined by the amplitude of the signal. As a consequence, the information carried by the largest-amplitude samples in the original signal is missing, the number of consecutive clipped samples may be large and these clipped intervals may occur frequently.

Audio declipping has been mainly addressed by linear prediction techniques [1–3]. In [3], declipping is addressed via a straightforward and basic usage of linear prediction: the autoregressive (AR) filter coefficients computed from the undistorted samples preceding clipping are used to predict the clipped samples. A more advanced approach has been proposed in [1] for the general problem of filling several gaps of consecutive missing samples simultaneously. While no explicit application is mentioned, the method is naturally appropriate for declipping. A single autoregressive model is considered for the region embedding the missing samples where the set of AR coefficients and the set of missing samples are alternately estimated by an Expectation Maximization-like iterative algorithm. Another approach based on linear prediction has been proposed in [2] where the autoregressive model is time-varying. In [4], the audio declipping problem was formulated as an l_2-norm minimization problem subject to a band-limited assumption, *i.e.* to the existence of zeros in the discrete Fourier transform of the original signal. Cubic interpolation has also been proposed for declipping in the *ClipFix* plug-in[1].

In this paper, we utilize Sparse Representation (SR) modeling of audio signals [6] – *i.e.* approximating audio frames by linear combinations of few atomic signals (columns of a dictionary matrix) – and leverage the image inpainting framework [7] where missing or masked pixel groups in an image are filled in. The audio declipping problem is formulated as an inverse problem, where one observes only a partial set consisting of reliable audio data – the un-clipped samples – while the remaining data to be estimated is treated as unknown. We employ an overlap-add (OLA) approach in conjunction with a constrained version of the Orthogonal Matching Pursuit (OMP) algorithm [8], to recover the sparse representation vectors of overlapping audio frames. The complete recovered signal is formed by filling in the missing audio samples in each frame and merging all frames in the OLA process.

The contributions of this paper are two-fold: 1) The formulation of the audio declipping problem as a SR recovery problem is, to the best knowledge of the authors, an original approach that enables the utilization of the rich theory and tools of SR modeling [9]. 2) A constrained OMP algorithm is introduced, which provides significantly improved results over its unconstrained (standard) version, by incorporating additional information inherent to the declipping problem. Performance evaluation over a collection of music and speech signals demonstrate superior results compared to existing methods.

This paper is organized as follows. The audio declipping problem is formulated in Sec. 2. The SR model and the constrained approach are detailed in Sec. 3. Experimental results are presented in Sec. 4. Conclusions are finally drawn in Sec. 5.

This work was supported by the EU Framework 7 FET-Open project FP7-ICT-225913-SMALL: Sparse Models, Algorithms and Learning for Large-Scale data and partly by the Israel Science Foundation grant no. 1031/08.

[1] See http://www.gaclrecords.org.uk/audacity.html

2. PROBLEM STATEMENT

We observe a clipped version $\mathbf{y} \in \mathbb{R}^L$ of an undistorted audio signal $\mathbf{s} \in \mathbb{R}^L$. The samples affected by clipping are located on the subset $\mathcal{I}^{\text{c}} \subset \mathcal{I}$ of the signal support $\mathcal{I} \triangleq \{1, 2, \cdots, L\}$, such that

$$\mathcal{I}^{\text{c}} \triangleq \{n \,|1 \leq n \leq L, |\mathbf{s}\,(n)| \geq \theta_{\text{clip}}\}\,, \tag{1}$$

where θ_{clip} is the clipping level. We thus consider the partition $\{\mathcal{I}^{\text{c}}, \mathcal{I}^{\text{r}}\}$ of the support $\mathcal{I} \triangleq \{1, 2, \cdots, L\}$, where $\mathcal{I}^{\text{r}} \triangleq \mathcal{I} \backslash \mathcal{I}^{\text{c}}$, such that the observed signal $\mathbf{y}$ is partitioned into the reliable (unclipped) $\mathbf{y}^{\text{r}}$ and clipped $\mathbf{y}^{\text{c}}$ portions as follows

$$\begin{cases} \mathbf{y}^{\text{r}} \triangleq \mathbf{y}\,(\mathcal{I}^{\text{r}}) = \mathbf{s}\,(\mathcal{I}^{\text{r}}) \\ \mathbf{y}^{\text{c}} \triangleq \mathbf{y}\,(\mathcal{I}^{\text{c}}) = \text{sign}\,(\mathbf{s}\,(\mathcal{I}^{\text{c}}))\,\theta_{\text{clip}}, \end{cases} \tag{2}$$

where sign $(\cdot)$ is the element-wise sign function. In matrix form, the observed reliable and clipped signal portions are given by

$$\begin{cases} \mathbf{y}^{\text{r}} = \mathbf{M}^{\text{r}}\mathbf{y} = \mathbf{M}^{\text{r}}\mathbf{s} \\ \mathbf{y}^{\text{c}} = \mathbf{M}^{\text{c}}\mathbf{y} = \mathbf{M}^{\text{c}}\,\text{sign}\,(\mathbf{s})\,\theta_{\text{clip}}, \end{cases} \tag{3}$$

where $\mathbf{M}^{\text{r}}$ is the reliable data measurement matrix obtained from the $L \times L$ identity matrix $\mathbf{I}_L$ by selecting the rows $\mathcal{I}^{\text{r}}$ associated with the reliable samples in $\mathbf{s}$. In a similar way, the clipped data measurement matrix $\mathbf{M}^{\text{c}}$ consists of the rows $\mathcal{I}^{\text{c}}$ in $\mathbf{I}_L$.

The audio declipping problem is an inverse problem, defined as the recovery of the original samples $\mathbf{s}$ from the observation $\mathbf{y}$. The detection of $\{\mathcal{I}^{\text{c}}, \mathcal{I}^{\text{r}}\}$ and the estimation $\hat{\theta}_{\text{clip}}$ of the clipping level θ_{clip} can be achieved by locating and selecting the maximum absolute value of the observed samples. We thus focus on the restoration of the clipped samples $\mathbf{s}\,(\mathcal{I}^{\text{c}})$ given $\mathbf{y}$, $\{\mathcal{I}^{\text{c}}, \mathcal{I}^{\text{r}}\}$ and θ_{clip}.

3. PROPOSED METHOD

We propose audio declipping algorithms for single channel waveforms. The proposed algorithms rely on frame-based processing, as detailed in Sec. 3.1, and emerge from SR modeling of audio signals, as presented in Sec. 3.2. A basic OMP algorithm is discussed in Sec. 3.3 and a constrained OMP algorithm is developed in Sec. 3.4.

3.1. Frame-based processing and reconstruction

Declipping is locally performed using a frame-by-frame processing. Every frame is independently restored and the full restored signal is formed utilizing an OLA approach [10]. We decompose the observed signal into overlapping frames $\mathbf{y}_i \in \mathbb{R}^N, N \ll L$, starting at time t_i, using a rectangular weighting window with length N: $\mathbf{y}_i\,(t) \triangleq \mathbf{y}\,(t + t_i)\,, 0 \leq t \leq N - 1$. By adapting the full signal problem statement to the local frames formulation, the reliable samples in the i-th frame are

$$\mathbf{y}_i^{\text{r}} = \mathbf{M}_i^{\text{r}}\mathbf{s}_i, \tag{4}$$

where $\mathbf{M}_i^{\text{r}}$ is the reliable data measurement matrix of the i-th frame obtained from $\mathbf{M}^{\text{r}}$ and $\mathbf{s}_i\,(t) \triangleq \mathbf{s}\,(t + t_i)$ is the i-th frame defined for $0 \leq t \leq N - 1$. We further define the supports $\{\mathcal{I}_i^{\text{c}}, \mathcal{I}_i^{\text{r}}\}$ of the clipped and reliable samples in the i-th frame, which can be simply computed from the full signal supports pair $\{\mathcal{I}^{\text{c}}, \mathcal{I}^{\text{r}}\}$. Once the estimation $\widehat{\mathbf{s}}_i$ of $\mathbf{s}_i$ is completed by any of the algorithms presented below, the reconstruction of the full signal is obtained as $\widehat{\mathbf{s}}(t) \triangleq \frac{\sum_i \mathbf{w}_s(t-t_i)\widehat{\mathbf{s}}_i(t-t_i)}{\sum_i \mathbf{w}_s(t-t_i)}$. In the proposed approaches, we utilize 64ms-frames with 75% overlap and sine windows for $\mathbf{w}_s$.

3.2. Sparse Representations modeling of audio frames

In the SR modeling framework [9], it is assumed that each frame is well approximated by a sparse linear combination of the columns of a given (possibly overcomplete) dictionary

$$\mathbf{s}_i \approx \mathbf{D}\mathbf{x}_i, \tag{5}$$

where $\mathbf{D} \in \mathbb{R}^{N \times K_D}$ is the dictionary, $N \leq K_D$, and $\mathbf{x}_i \in \mathbb{R}^{K_D \times 1}$ is a sparse vector: $\|\mathbf{x}_i\|_0 \ll L$, where the l_0 norm[2] $\|\mathbf{x}_i\|_0$ counts the non-zero components in $\mathbf{x}_i$. As a consequence, we can also utilize the SR model for the observed reliable samples in each frame

$$\mathbf{y}_i^{\text{r}} \triangleq \mathbf{M}_i^{\text{r}}\mathbf{s}_i \approx \mathbf{M}_i^{\text{r}}\mathbf{D}\mathbf{x}_i. \tag{6}$$

We propose to estimate the unknown samples $\mathbf{s}_i\,(\mathcal{I}_i^{\text{c}})$ by recovering the SR vector of each frame $\mathbf{x}_i$, given only the reliable observed samples $\mathbf{y}_i^{\text{r}}$, the support partition $\{\mathcal{I}_i^{\text{c}}, \mathcal{I}_i^{\text{r}}\}$ and the estimated clipping level $\hat{\theta}_{\text{clip}}$. Once the SR vector is recovered as $\hat{\mathbf{x}}_i$, the unknown samples are estimated according to

$$\hat{\mathbf{s}}_i\,(\mathcal{I}_i^{\text{c}}) \approx \mathbf{M}_i^{\text{c}}\mathbf{D}\hat{\mathbf{x}}_i, \tag{7}$$

where $\mathbf{M}_i^{\text{c}}$ is the clipped data measurement matrix of the i-th frame obtained from $\mathbf{M}^{\text{c}}$. In the following we overview two approaches to solve this problem, based on approximate solutions to the l_0 norm minimization problem.

3.3. A basic Matching Pursuit algorithm for audio declipping

The proposed approaches seek for the sparsest representation of each audio frame, by approximating a solution to the following optimization problem

$$\hat{\mathbf{x}}_i = \arg\min_{\mathbf{x}_i} \|\mathbf{x}_i\|_0 \;\text{ s.t. }\; \|\mathbf{y}_i^{\text{r}} - \mathbf{M}_i^{\text{r}}\mathbf{D}\mathbf{x}_i\|_2^2 \leq \theta_i^{\epsilon}. \tag{8}$$

A direct solution of (8) is infeasible since the l_0 norm leads to an NP-hard problem. An approximate solution is given by applying the OMP algorithm [8], which successively approximates the sparsest solution. The inpainting OMP algorithm [7], detailed in Table 1, is a slightly modified version of the classical OMP algorithm [8] in the sense that all dictionary columns are internally re-normalized to unit norm due to the availability of only the reliable samples. The algorithm stops iterating as soon as either the number of non-zero components exceeds the maximum sparsity level K_{max}, or the residual energy drops below the threshold θ_i^{ϵ}.

3.4. Algorithmic enhancements for audio declipping

Recovering clipped signals can be performed with the algorithm presented in Sec. 3.3, by treating the clipped samples as completely unknown. However, side information inherent to this problem can be integrated as additional constraints into equation (8). Let $\mathbf{M}_i^{\text{c+}}$ (resp. $\mathbf{M}_i^{\text{c-}}$) be the matrix such that $\mathbf{M}_i^{\text{c+}}\mathbf{s}_i$ (resp. $\mathbf{M}_i^{\text{c-}}\mathbf{s}_i$) is the vector of positive (resp. negative) clipped samples. The matrices $\mathbf{M}_i^{\text{c+}}$ and $\mathbf{M}_i^{\text{c-}}$ are known according to the sign of each clipped sample, and the following set of of constraints can be defined for the set of missing samples[3]

[2]Note that the l_0 norm is not a standard norm as it does not obey $\|\alpha x\|_0 = \alpha\,\|x\|_0$ for any positive scalar α, however, the term "norm" is traditionally associated with this quantity.

[3]Inequalities are defined element-wise for notation convenience.

Table 1. OMP Inpainting Algorithm

Input: $\mathbf{y}_i^\mathrm{r},\ \mathbf{M}_i^\mathrm{r},\ \mathbf{D},\ K_{\mathbf{max}},\ \theta_i^\epsilon$
Initialize : • Dictionary $\widetilde{\mathbf{D}} = \left[\widetilde{\mathbf{d}}_1, \ldots, \widetilde{\mathbf{d}}_{K_D}\right] = \mathbf{M}_i^\mathrm{r} \times \mathbf{D} \times \mathbf{W}$, where $\mathbf{W}$ is a diagonal matrix such that diagonal component j equals the inverse of the norm of column j of the matrix $\mathbf{M}_i^\mathrm{r} \times \mathbf{D}$. • Iteration counter $k = 0$ • Support set $\Omega_0 = \emptyset$ • Residual $\mathbf{r_0} = \mathbf{y_i^r}$
Repeat until $\mathbf{k} = \mathbf{K}_{\max}$ **OR** $\|\mathbf{r_k}\|_\mathbf{2}^\mathbf{2} < \theta_\mathbf{i}^\epsilon$ • Increment iteration counter $k = k + 1$ • Select atom: find $j = \arg\max_j \mid <\mathbf{r_{k-1}}, \widetilde{\mathbf{d}}_\mathbf{j}> \mid$ • Update Support $\Omega_k = \Omega_{k-1} \cup j$ • Update current solution $\mathbf{x_k} = \arg\min_\mathbf{u} \|\mathbf{y_i^r} - \widetilde{\mathbf{D}}_{\mathbf{\Omega_k}}\mathbf{u}\|_\mathbf{2}$ • Update Residual $\mathbf{r_k} = \mathbf{y_i^r} - \widetilde{\mathbf{D}}_{\mathbf{\Omega_k}}\mathbf{x_k}$
Output: $\mathbf{x_i} = \mathbf{W}\mathbf{x_k}$

Table 2. Constrained OMP Declipping Algorithm

Input: $\mathbf{y}_i^\mathrm{r},\ \mathbf{M}_i^\mathrm{r},\ \mathbf{D},\ K_{\mathbf{max}},\ \theta_i^\epsilon,\ \hat{\theta}_{\mathbf{clip}},\ \hat{\theta}_{\mathbf{max}}$
Initialize : • $\widetilde{\mathbf{D}} = \mathbf{M}_i^\mathrm{r} \times \mathbf{D} \times \mathbf{W}$ • $k = 0, \Omega_0 = \emptyset, \mathbf{r_0} = \mathbf{y_i^r}$
Repeat until $\mathbf{k} = \mathbf{K}_{\max}$ **OR** $\|\mathbf{r_k}\|_\mathbf{2}^\mathbf{2} < \theta_\mathbf{i}^\epsilon$ • Increment iteration counter $k = k + 1$ • Select atom: find $j = \arg\max_j \mid <\mathbf{r_{k-1}}, \widetilde{\mathbf{d}}_\mathbf{j}> \mid$ • Update Support $\Omega_k = \Omega_{k-1} \cup j$ • Update current solution $\mathbf{x_k} = \arg\min_\mathbf{u} \|\mathbf{y_i^r} - \widetilde{\mathbf{D}}_{\mathbf{\Omega_k}}\mathbf{u}\|_\mathbf{2}$ • Update Residual $\mathbf{r_k} = \mathbf{y_i^r} - \widetilde{\mathbf{D}}_{\mathbf{\Omega_k}}\mathbf{x_k}$
Final solution update :$\mathbf{x_k} = \arg\min_\mathbf{u} \|\mathbf{y_i^r} - \widetilde{\mathbf{D}}_{\mathbf{\Omega_k}}\mathbf{u}\|_\mathbf{2}$ s.t. $\begin{cases} \hat{\theta}_{\max} \geq \mathbf{M}_i^{\mathrm{c+}}\mathbf{DWu} \geq \hat{\theta}_{\mathrm{clip}} \\ -\hat{\theta}_{\max} \leq \mathbf{M}_i^{\mathrm{c-}}\mathbf{DWu} \leq -\hat{\theta}_{\mathrm{clip}} \end{cases}$
Output: $\mathbf{x_i} = \mathbf{W}\mathbf{x_k}$

$$\begin{cases} \mathbf{M}_i^{\mathrm{c+}}\mathbf{s}_i \geq \hat{\theta}_{\mathrm{clip}} \\ \mathbf{M}_i^{\mathrm{c-}}\mathbf{s}_i \leq -\hat{\theta}_{\mathrm{clip}}. \end{cases} \tag{9}$$

This set of constraints can be further augmented by introducing an upper limit on the absolute value of the recovered samples $\hat{\theta}_{\max}$ as

$$\begin{cases} \mathbf{M}_i^{\mathrm{c+}}\mathbf{s}_i \leq \hat{\theta}_{\max} \\ \mathbf{M}_i^{\mathrm{c-}}\mathbf{s}_i \geq -\hat{\theta}_{\max}. \end{cases} \tag{10}$$

The upper limit $\hat{\theta}_{\max}$ is an optional parameter, that can be roughly approximated as $\hat{\theta}_{\max} \triangleq Q \times \hat{\theta}_{\mathrm{clip}}$ for some positive scalar Q.

The declipping version of the l_0 norm minimization problem (8) is given by

$$\widehat{\mathbf{x}_i} = \arg\min_{\mathbf{x}_i} \|\mathbf{x}_i\|_0 \text{ s.t. } \begin{cases} \|\mathbf{y}_i^\mathrm{r} - \mathbf{M}_i^\mathrm{r}\mathbf{D}\mathbf{x}_i\|_2^2 \leq \theta_i^\epsilon \\ \hat{\theta}_{\max} \geq \mathbf{M}_i^{\mathrm{c+}}\mathbf{D}\mathbf{x}_i \geq \hat{\theta}_{\mathrm{clip}} \\ -\hat{\theta}_{\max} \leq \mathbf{M}_i^{\mathrm{c-}}\mathbf{D}\mathbf{x}_i \leq -\hat{\theta}_{\mathrm{clip}} \end{cases} \tag{11}$$

We propose to approximate the solution of (11) by incorporating the constraints (9) and (10) into the final solution update stage of the OMP Inpainting algorithm, as presented in Table 2. In the following, the algorithm including (9) only and the one including both (9) and (10) will be referred to as the single-constraint OMP algorithm and the dual-constraint OMP algorithm respectively.

4. EXPERIMENTAL RESULTS

The experiments are conducted on a dataset of ten phone-quality speech signals sampled at 8 kHz and a dataset of ten music signals sampled at 16 kHz (*i.e.* with higher quality than phone speech). Each test signal is 5-second long and is part of the freely available material of the 2008 Signal Separation Evaluation Campaign [4]. The test data shows a large diversity of audio mixtures and isolated sources, including male and female speech from different speakers, singing voice, pitched and percussive musical instruments. Each original signal is normalized so that the maximum amplitude is 1. Each sound is then artificially clipped with successive clipping levels, from 0.2 up to 0.9 with a 0.1-step.

In the proposed algorithms, an overcomplete Discrete Cosine Transform (DCT) dictionary was used. This choice is motivated by the wide usage of DCT to code or estimate audio signals [6]. A specific training dataset was used to tune the parameters of the inpainting algorithms. The following values were obtained: the frame length is set to 64 ms (*i.e.* $N \triangleq 512$ and $N \triangleq 1024$ samples at 8 kHz and 16 kHz respectively); a 75% frame overlap is used; the number of atoms – columns – in the overcomplete DCT dictionary is set to twice the number of samples in a frame; fixed values are set to the stopping criteria of the OMP algorithm: $K_{\max} \triangleq \frac{N}{4}$ and $\theta_i^\epsilon \triangleq \theta_\epsilon \times \#\mathcal{I}_i^\mathrm{r}$, where $\theta_\epsilon \triangleq 10^{-6}$ is a fixed parameter and $\#\mathcal{I}_i^\mathrm{r}$ is the number of reliable samples in the ith frame. The dual constraint was used with $Q = 4$.

For comparison purposes, we implemented the method [1] by Janssen et al. based on linear prediction. In each frame, a single AR model is considered for both the reliable observed samples and the latent missing samples, which are estimated in an iterative algorithm. The AR order was set to $3m + 2$, as recommended by the authors, where m is the number of missing samples. The *ClipFix* plug-in based on cubic interpolation (see Sec.1) was also tested.

Clipping restoration is illustrated in Fig. 2 when the clipping level is 0.2. Here, the dual-constraint OMP algorithm is applied to an example of music signal, where one can observe that the recon-

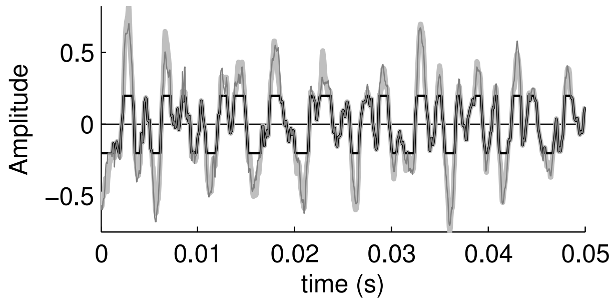

Fig. 2. Restoration of a music signal by the dual-constraint algorithm: original (light gray), clipped (black), estimate (dark gray).

[4]http://sisec2008.wiki.irisa.fr/

structed samples are close to the original signal.

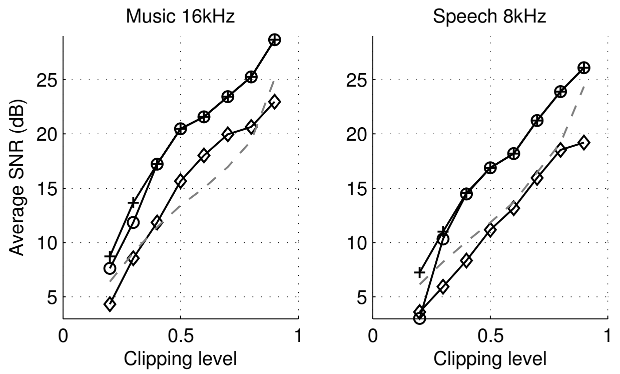

Fig. 3. Performance of the proposed algorithms, as a function of the clipping level: the SNR is averaged for the music (left) and the speech (right) datasets. Curve specification: unconstrained OMP ("◊"); single ("o") and dual ("+") constraint OMP; initial clipped signal (dashed gray).

The performance on the overall datasets is assessed by the signal-to-noise ratio (SNR) computed on the clipped samples

$$\mathrm{SNR}_c\left(\mathbf{s}, \widehat{\mathbf{s}}\right) \triangleq 10 \log \frac{\left\|\mathbf{s}\left(\mathcal{I}^c\right)\right\|_2^2}{\left\|\mathbf{s}\left(\mathcal{I}^c\right)-\widehat{\mathbf{s}}\left(\mathcal{I}^c\right)\right\|_2^2} \tag{12}$$

which reflects the reconstruction performance per estimated sample and differs from the SNR computed on the full signals by an offset $10 \log \frac{\|\mathbf{s}\|_2^2}{\|\mathbf{s}(\mathcal{I}^c)\|_2^2}$ that does not depend on the declipping algorithm.

The performance of the proposed algorithms are reported in Fig. 3. The single-constraint and dual-constraint algorithms enhance the SNR by 4 dB and 4.5 dB on the average, respectively. For almost all test sounds, they significantly improve the unconstrained OMP algorithm: indeed, the latter algorithm happens to reach poor results, even degrading the distorted signal in the case of the speech dataset. This shows that methods based on SR, if efficient under random-measurement conditions [7], cannot straightforwardly recover partially-sampled signals when groups of missing samples are involved. The dual-constraint OMP algorithm reaches better results than the single-constraint algorithm when the clipping level is about $0.2 - 0.3$. This corresponds to the range where the approximate value $\hat{\theta}_{\max}$ is close to the actual maximum value as well as to the most degraded signals. A close analysis of the individual restored sounds reveals that large spikes are avoided thanks to the maximum value constraint. In a practical application, the maximum value $\hat{\theta}_{\max}$ should be adjusted by the user until the best audio quality is achieved. In Fig. 4 the dual-constraint algorithm is compared against existing methods. It outperforms Janssen's method by 1.9 dB on the average. The *ClipFix* plug-in reaches poor results, below all the reported ones.

5. CONCLUSIONS

We presented a novel sparse representation based approach for the restoration of clipped audio signals. In the proposed approach, the sign pattern of the clipped samples and their maximum absolute value are integrated into a constrained OMP algorithm. Performance

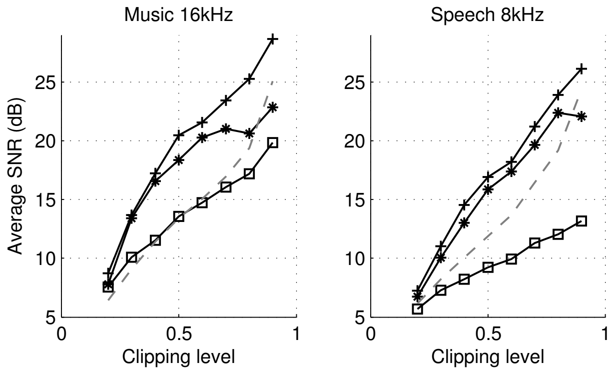

Fig. 4. Comparison with existing methods: Dual constraint OMP ("+"); Janssen's approach [1] ("∗"); *ClipFix* plug-in ("□"); initial clipped signal (dashed gray).

evaluation with a relatively simple dictionary - an overcomplete DCT - demonstrated an advantage compared to existing methods and the unconstrained OMP. In future research our approach could be adapted to address other audio restoration problems such as click removal and packet loss concealment.

6. REFERENCES

[1] A. Janssen, R. Veldhuis, and L. Vries, "Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes," *IEEE Trans. Acoustics, Speech and Sig. Proc.*, vol. 34, no. 2, pp. 317 – 330, apr 1986.

[2] S. J. Godsill, P. J. Wolfe, and W. N. W. Fong, "Statistical model-based approaches to audio restoration and analysis," *Journal of New Music Research*, vol. 30, no. 4, 2001.

[3] A. Dahimene, M. Noureddine, and A. Azrar, "A simple algorithm for the restoration of clipped speech signal," *Informatica*, vol. 32, pp. 183–188, 2008.

[4] J. S. Abel and J. O. Smith, "Restoring a clipped signal," in *Proc. of ICASSP*, Toronto, Canada, May 1991.

[5] S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration - A Statistical Model-based Approach*, Springer-Verlag, 1998.

[6] M. D. Plumbley, T. Blumensath, L. Daudet, R. Gribonval, and M. Davies, "Sparse representations in audio and music: from coding to source separation," *Proc. of the IEEE*, vol. 98, no. 6, 2010.

[7] M. Elad, *Sparse and Redundant Representations - From Theory to Applications in Signal and Image Processing*, Springer New-York, 2010.

[8] Y.C. Pati, R. Rezaiifar, and P.S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Nov. 1993, pp. 40 –44 vol.1.

[9] A.M. Bruckstein, D.L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM review*, vol. 51, no. 1, pp. 34–81, 2009.

[10] D. Griffin and Jae Lim, "Signal estimation from modified short-time fourier transform," *IEEE Trans. Acoustics, Speech and Sig. Proc.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.